LA-UR-12-1352

Approved for public release;
distribution is unlimited.

Title: Subband coding for large-scale scientific simulation data using JPEG 2000

Author(s): Christopher M. Brislawn, Jonathan L. Woodring, Susan M. Mniszewski, David E. DeMarle, and James P. Ahrens

Intended for: Southwest Symposium on Image Analysis & Interpretation
IEEE Computer Society
Santa Fe, NM, April 22-24, 2012

Los Alamos
NATIONAL LABORATORY
EST. 1943

Subband Coding for Large-Scale Scientific Simulation Data Using JPEG 2000

Christopher M. Brislawn, Jonathan L. Woodring,
Susan M. Mniszewski, James P. Ahrens
(Los Alamos National Laboratory)

David E. DeMarle (Kitware, Inc.)

Southwest Symposium on Image Analysis & Interpretation
Santa Fe, NM, April 2012

Scientific Simulation Data

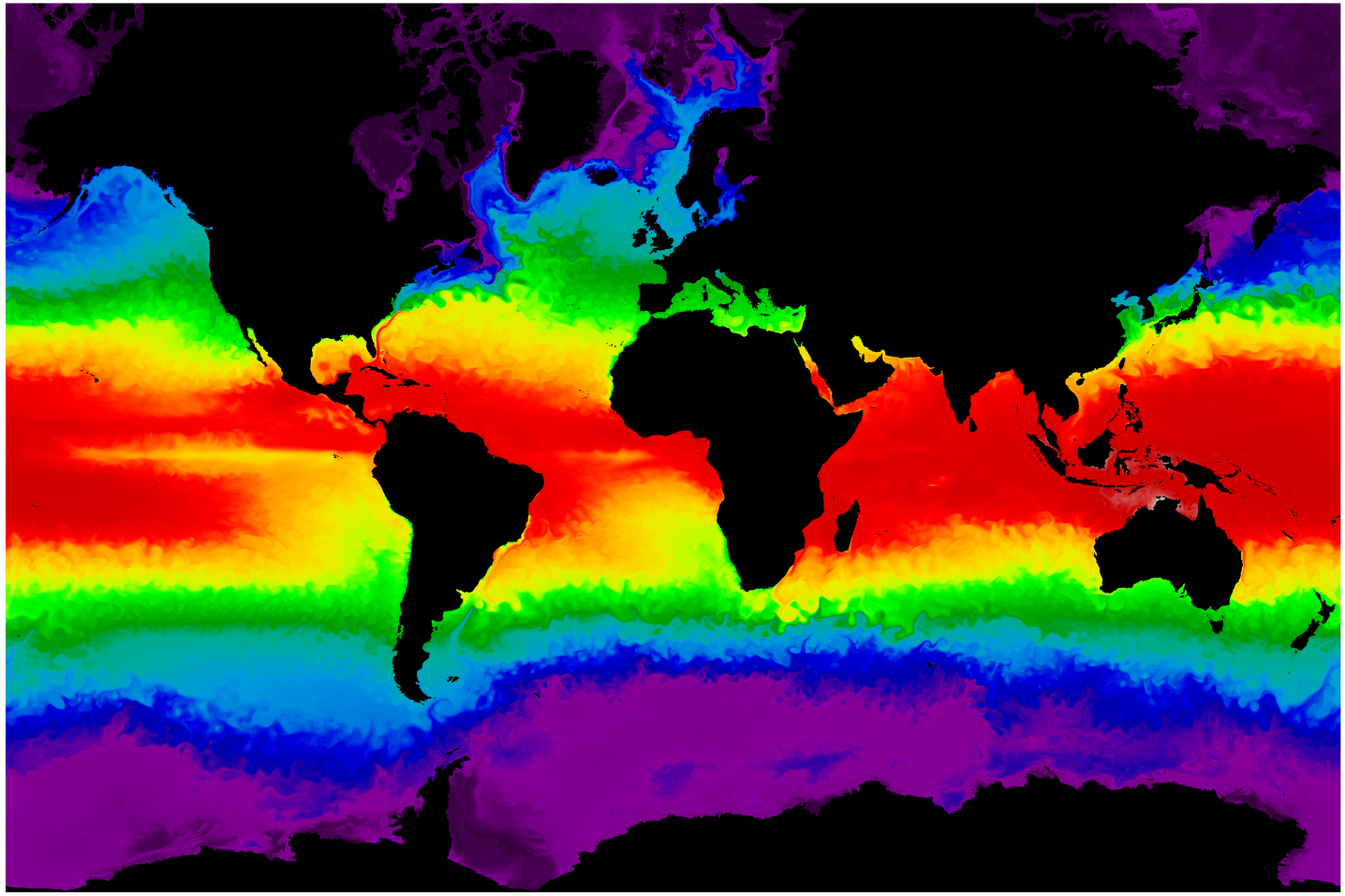
Los Alamos NATIONAL LABORATORY EST. 1943

- Distinct and rather complicated niche in data management and visualization
  - Floating point, analyzed/visualized at 32 bits/sample, fills its numerical range continuously
  - (Often) 2-3 space dimensions + time, with some smoothness
  - Often multicomponent data (e.g., multiple physical fields)
  - May be uniformly gridded, nonuniform but structured, or completely unstructured; stationary or time-varying
- The common feature is size: big and getting bigger

Example:
Global ocean circulation model

Los Alamos NATIONAL LABORATORY EST. 1943

- Parallel Ocean Program (POP) model, run @ ORNL
  - 3600 x 2400 x 42 grid (1/10 degree), 11 km nominal resolution
  - Temp., salinity, x-vel., y-vel. at 32 bits/sample = 5.4 GB/timestep



A sampling of petascale supercomputers (10^{15} ops/sec.)

Los Alamos NATIONAL LABORATORY EST. 1943

- IBM Roadrunner (2008) @ Los Alamos
  - 122,400 cores: hybrid Opteron & PowerXCell architecture; 106 TB memory
  - 1.04 petaflops/s (PF) on Linpack benchmark; power consumption: 2.4 MW
- Cray Jaguar (2008) @ Oak Ridge
  - 224,256 Opteron cores; 300 TB memory; 1.76 PF (Linpack); 7.0 MW
- NUDT Tianhe-1A (2010) @ Tianjin, China
  - 186,368 Xeon & NVIDIA cores; 229 TB; 2.57 PF (Linpack); 4.0 MW
- Fujitsu K Computer (2011) @ Kobe, Japan
  - 705,024 SPARC64 cores; 1410 TB memory; 10.51 PF (Linpack); 12.7 MW
- Unfortunately, petascale computing is still inadequate for resolving some important phenomena:
  - natural length scales (~1 km) of critical features in global ocean and atmospheric models

Petascale inadequate for:

Los Alamos NATIONAL LABORATORY EST. 1943

- inherent uncertainties in turbulent CFD
  - turbulence-chemistry interactions in combustion modeling
  - 85% of world's power supply is generated by burning fossil fuels
- bridging atomistic-molecular-mesoscale-macroscopic scales in materials science and biological systems
- 3-D neutron transport in nuclear reactor cores
- plasma and materials problems in fusion energy production

- Thus, the US Dept. of Energy (DOE) is committed to developing *exascale* computing (10^{18} ops/sec.)
- Sobering expectations of exascale computing:
  - Transistors still getting smaller per Moore's law, but energy efficiency not improving enough for clock freq's to keep increasing
  - Clock stagnation at 2-3 GHz forcing massive parallelism
    - 1 exaflop / 1 GHz = 10^9 cores
  - Memory budget of 100 PB is just 100 MB/core
  - DOE decided a power budget of *just 20 MW* is all they can afford for exascale computing!

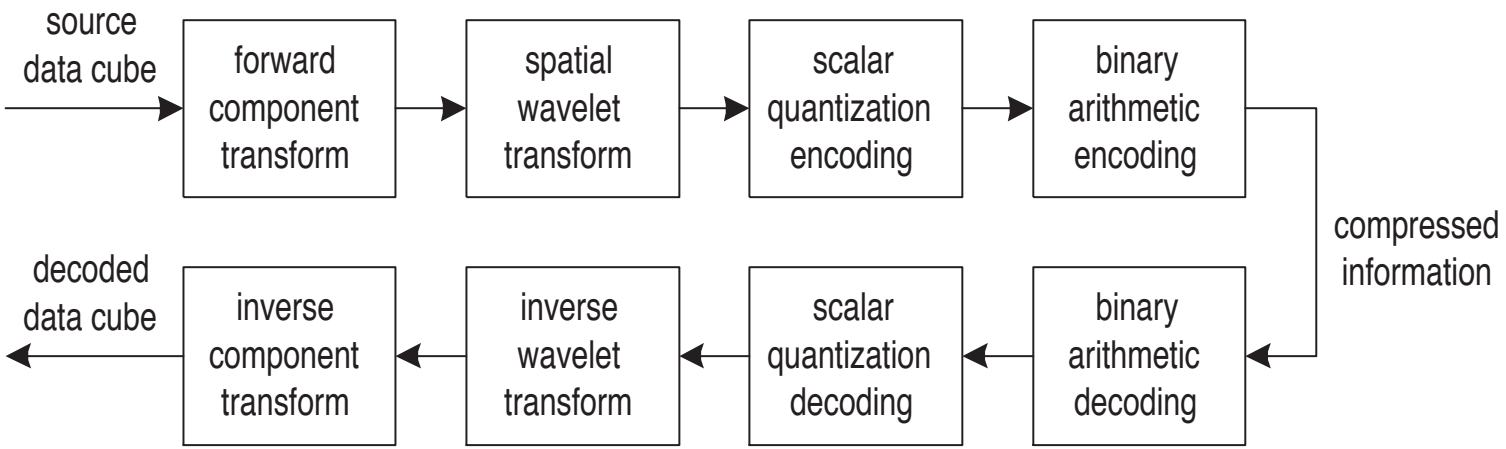
Anticipated realities of exascale

Los Alamos NATIONAL LABORATORY EST. 1943

- *Data movement* is single biggest energy cost
  - Energy to fetch a value from (non-cache) memory will exceed energy of a floating point operation
- Current thinking calls for around 1 exabyte of disk storage (1000 PB, or just 10x core memory)
- I/O bandwidth to storage estimated at ≤ 60 TB/s
  - 100 PB memory / 60 TB/s ≥ 28 minutes to write to disk
- We *cannot save* more than a tiny fraction of the data
  - Much analysis and visualization will have to be done *in situ* (while data is still in the machine)
- **Conclusion**: Can't afford *not* to use the most bandwidth-efficient communications coding available for exascale data
  - Computational cost of source coding is *not* the issue: can do a *lot* more flops in exchange for a significant reduction in data

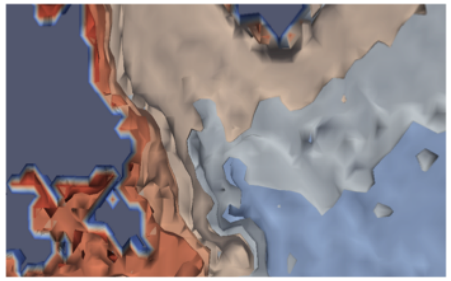
Enter JPEG 2000

- Family of international standards for source coding (compression) of digital imagery (ISO/IEC 15444-x)
  - Supports huge array dimensions, high bit depths, 3-D
  - Highly scalable w.r.t. spatial resolution & reconstructed sample precision, random access to regions-of-interest
  - Standardized HTTP client-server protocol for interactive browsing and retrieval (JPEG 2000 Part 9---"JPIP")
  - "Subband coding scheme": discrete wavelet transforms, embedded binary arithmetic bit-plane coding



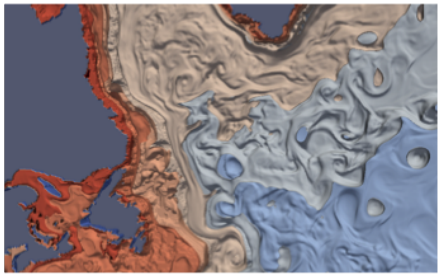
Software Environment

- Currently using “Kakadu” JPEG 2000 software with our own ParaView JPEG 2000 reader + prototype Paraview adaptive multiresolution (AMR) interface:

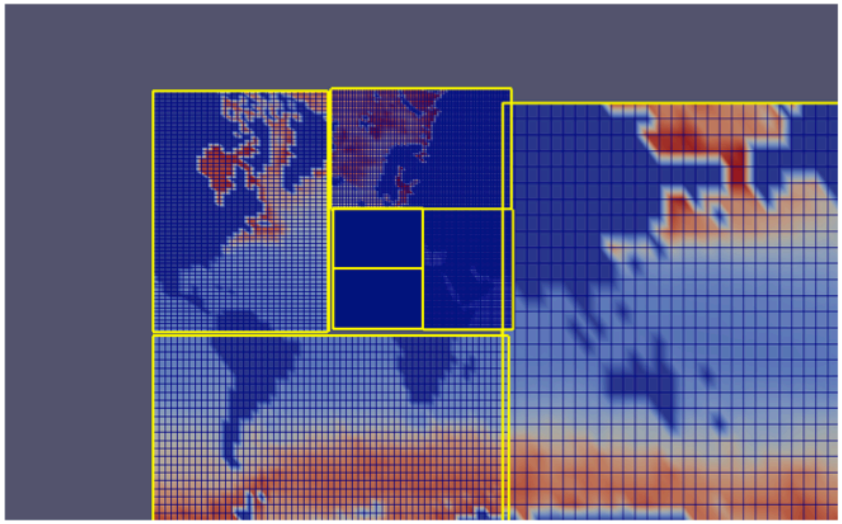


Low-resolution ROI

Streaming interactive update



High-res refinement



Qualitative performance with isocontouring on POP salinity data

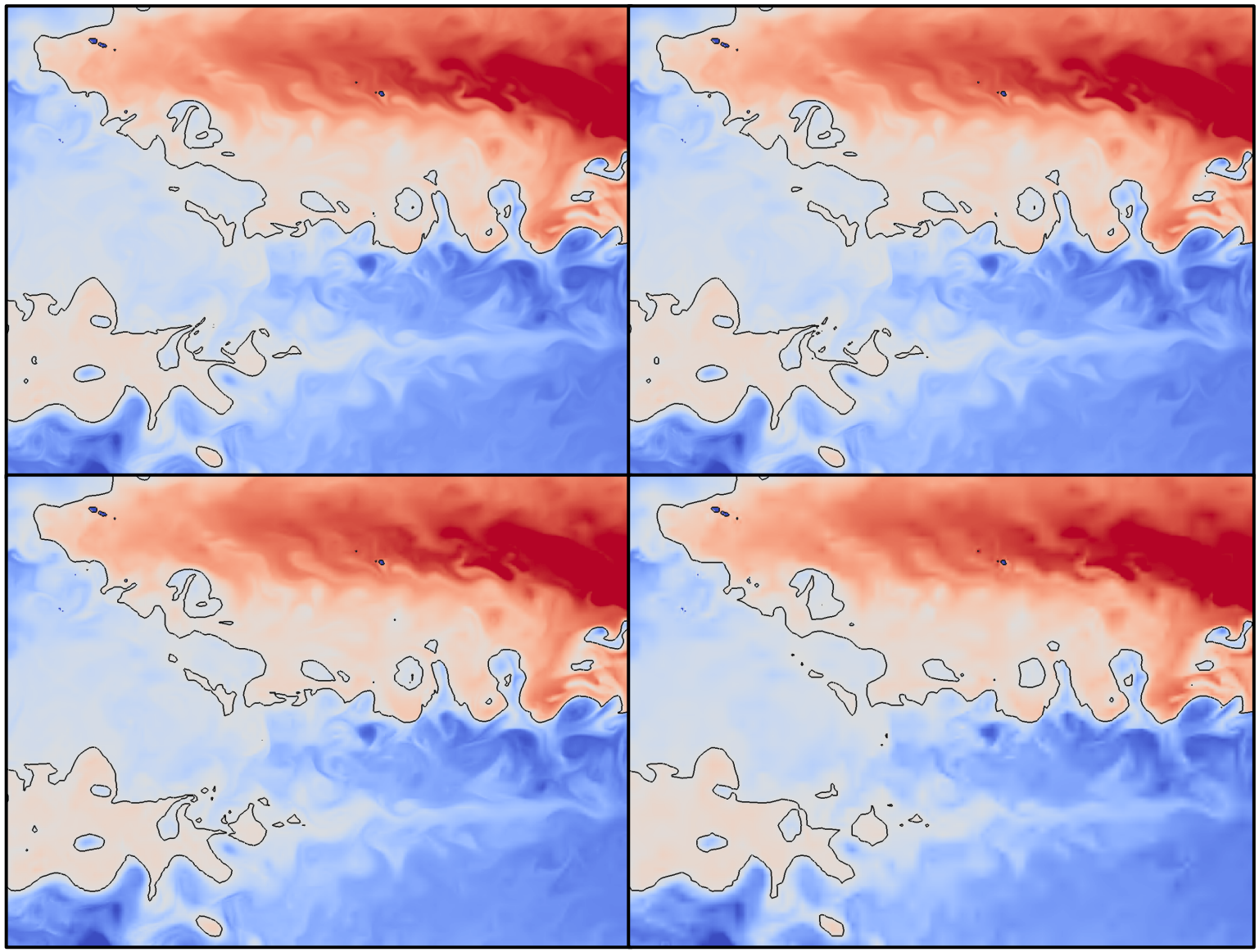
475x358 detail reconstructed at:

(T-L) 8.0 bpp, max.err = 1.5×10^{-9}

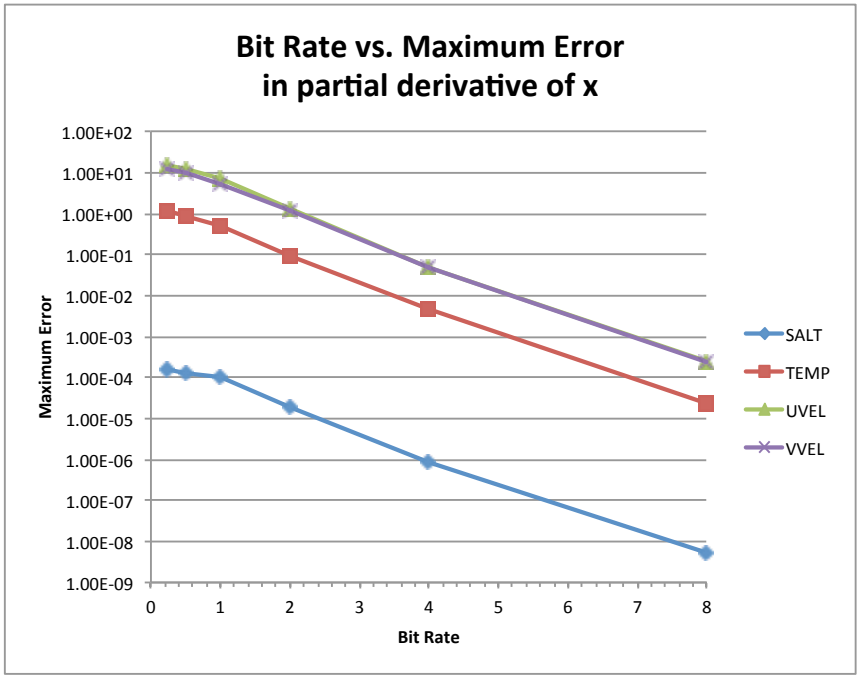
(T-R) 1.0 bpp, max.err = 2.3×10^{-5}

(B-L) 0.5 bpp, max.err = 8.6×10^{-5}

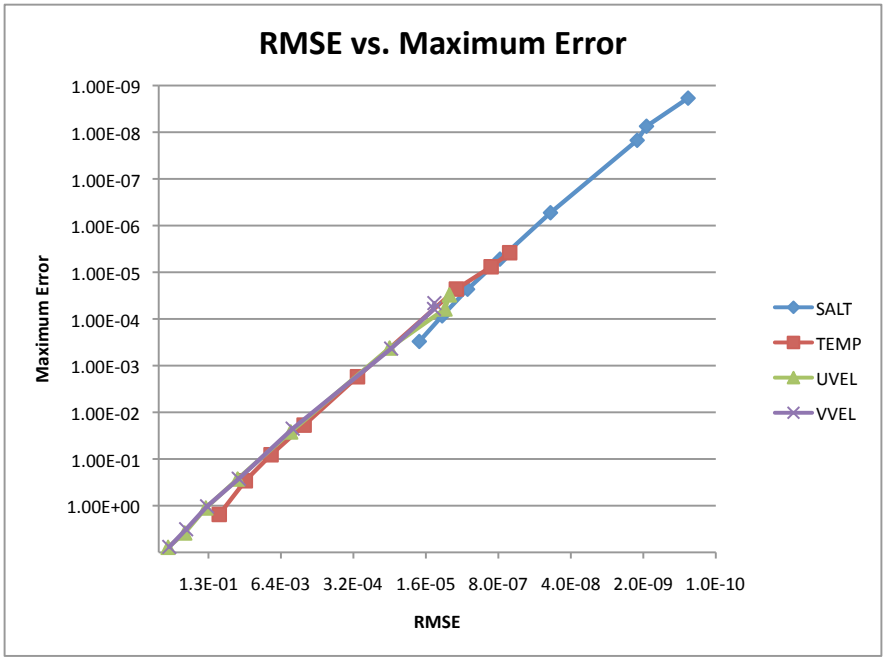
(B-R) 0.25 bpp, max.err = 3.0×10^{-4}



Quantitative rate-distortion performance metrics



Admirably linear behavior, esp. for max. error in numerical derivatives



Empirically, log max. error is proportional to log RMSE, with a data-*independent* constant of proportionality,

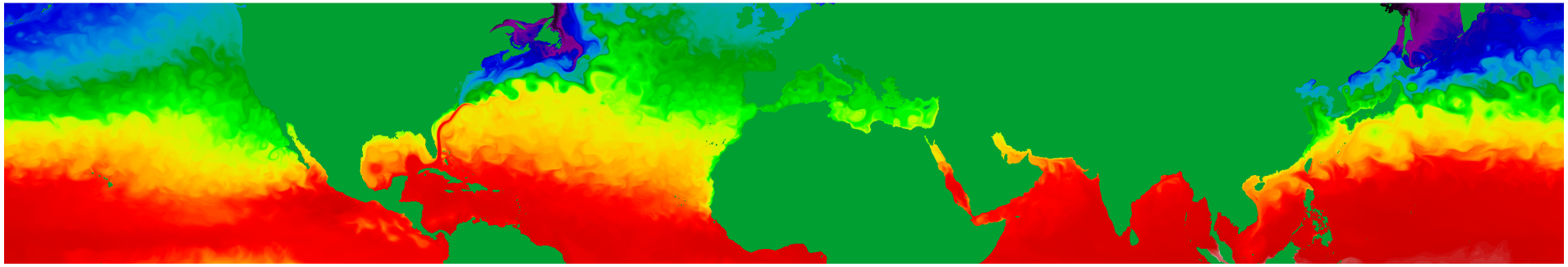
$$\log(\text{max.err}) \approx 10 * \log(\text{RMSE})$$

$$\text{max.err} \approx (\text{RMSE})^{10}$$

(incorrectly stated in paper)

What about masked regions?

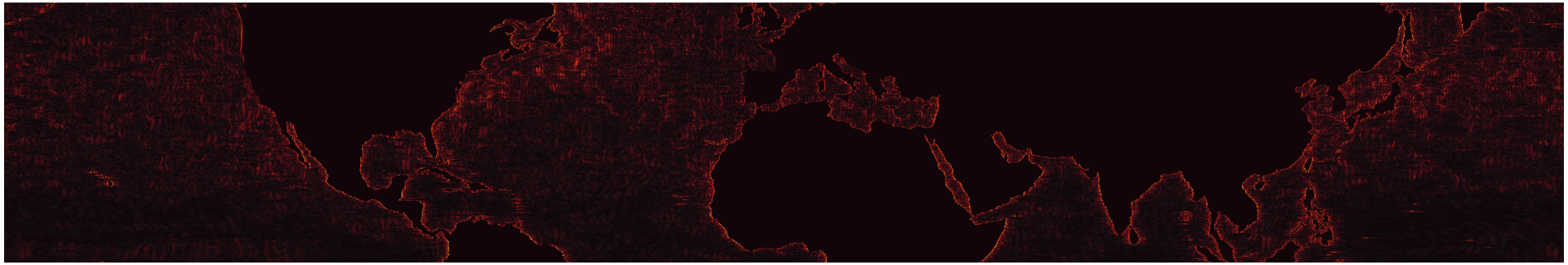
- Fill with a constant value like the global mean of the data.
  - Cheap, hands-off, compressible, but it leaves discontinuities that increase signal entropy and cause artifacts when data is reconstructed at reduced rates.



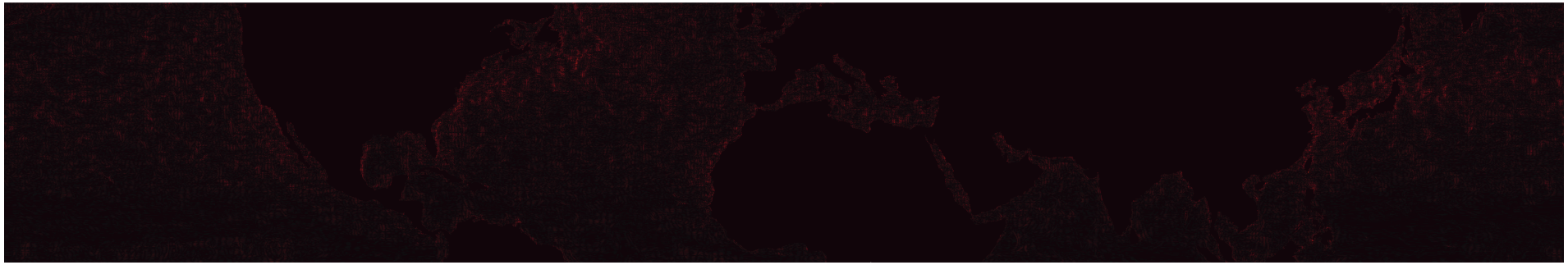
POP surface temperature (25% of the global array), continents filled with global mean of ocean data

Difference Images @ 0.5 bits/sample

- Land masses masked off, error images scaled to have identical “hot” color maps.
- Error image for global mean-filled interpolation. Note concentration of large errors near coastlines:

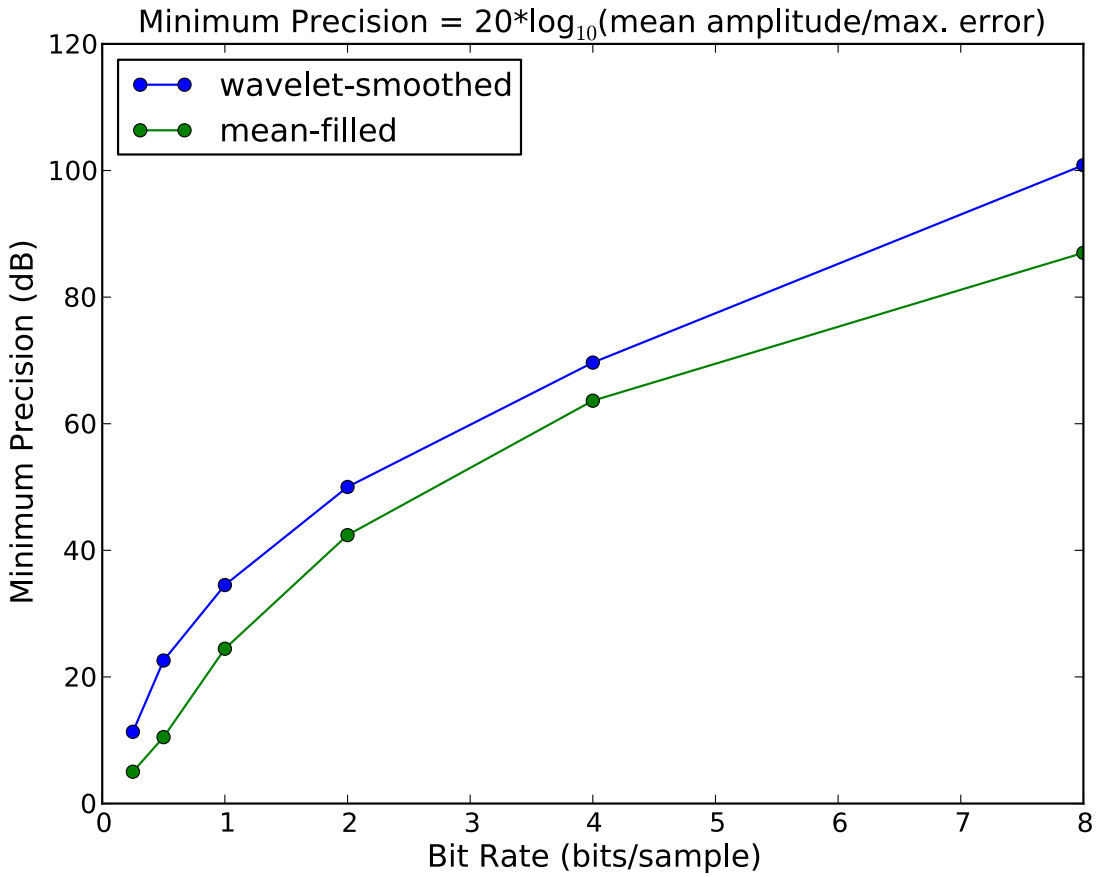


- Error image for wavelet-smoothed interpolation:



Minimum Precision vs. Rate

- Error taken over initialized (meaningful) data only.
  - Difference of 6 to 14 dB: 1 to 2.3 bits diff. in max error.



Conclusions

- Next steps:
  - Enable floating point I/O, entropy estimation for JPEG 2000
  - Extend to 3-D using JPEG 2000 Part 10
  - Interactive client-server semantics between Paraview clients and JPIP servers
  - 2-D and 3-D wavelet-smoothed interpolation of masked regions using matrix-free conjugate-gradient solver
  - In-situ JPEG 2000 encoding for exascale simulations

- Funded by the DOE Office of Science Program in Advanced Scientific Computing Research (ASCR)
  - program manager: Dr. Lucy Nowell